Holonomic gradient method for the distribution function of the largest root of a Wishart matrix

Hiroki Hashiguchi*, Yasuhide Numata†‡, Nobuki Takayama§‡ and Akimichi Takemura†‡

January 2012

Abstract

We apply the holonomic gradient method introduced by Nakayama et al. [19] to the evaluation of the exact distribution function of the largest root of a Wishart matrix, which involves a hypergeometric function ${}_1F_1$ of a matrix argument. Numerical evaluation of the hypergeometric function has been one of the longstanding problems in multivariate distribution theory. The holonomic gradient method offers a totally new approach, which is complementary to the infinite series expansion around the origin in terms of zonal polynomials. It allows us to move away from the origin by the use of partial differential equations satisfied by the hypergeometric function. From numerical viewpoint we show that the method works well up to dimension 10. From theoretical viewpoint the method offers many challenging problems both to statistics and D-module theory.

Keywords and phrases: D-modules, Gröbner basis, hypergeometric function of a matrix argument, zonal polynomial

1 Introduction

For multivariate distribution theory in statistics, the theory of zonal polynomials and hypergeometric functions of matrix arguments, introduced by A.T. James and other authors, was a very important development in the 1950's. They allowed explicit expressions of density functions and cumulative distribution functions of basic test statistics under non-null cases. Zonal polynomials are based on the representation theory of real general linear group and they possess many interesting combinatorial properties. Properties and

*Graduate School of Science and Engineering, Saitama University

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo

‡JST CREST

§Department of Mathematics, Kobe University

applications of zonal polynomials and hypergeometric functions of matrix arguments are surveyed in Gross and Richards [4] and Richards [21]. Zonal polynomials are special cases of Jack polynomials, whose properties have been intensively studied by many mathematicians. See for example Chapters VI and VII of Macdonald [14] and Stanley [25]. Jack polynomials are further generalized to Macdonald polynomials (see, e.g., Kuznetsov and Sahi [13]).

Despite the above nice mathematical properties of zonal polynomials and hypergeometric functions of matrix arguments, from practical viewpoint they were not really useful for computations. Coefficients of zonal polynomials can be computed only through nontrivial combinatorial recursions. Although very ingenious recursion algorithms have been recently developed (Koev and Edelman [10]), computing zonal polynomials of large degrees remains to be a difficult problem because of inherent combinatorial complexities. Also, the convergence of infinite series expansion of hypergeometric functions of a matrix argument in terms of zonal polynomials was found to be slow (Muirhead [17], Hashiguchi and Niki [5]). Since the expansion of the hypergeometric function in terms of zonal polynomials is the expansion *at the origin*, the convergence for large values of the argument is necessarily slow.

The holonomic gradient method allows us to move away from the origin by the use of partial differential equations. Thus our approach provides a promising new method for attacking a longstanding problem in multivariate statistics. Our holonomic gradient method is, in spirit, on the track of the holonomic systems approach to combinatorial identities by Zeilberger [32]. Note that the series expansion and our holonomic gradient method are in fact complementary methods, because our method needs the series expansion for obtaining initial values for the partial differential equations.

The main purpose of this paper is to verify the performance of holonomic gradient method for ${}_1F_1$. We found that a straightforward implementation of the holonomic gradient method works well for dimensions up to 10.

Butler and Wood [2] showed that the Laplace method gives a very good approximation to ${}_1F_1$ even for a high dimension, e.g., $m = 32$. However the Laplace method needs a peaked density function, which corresponds to a large degrees of freedom. Our method is an exact method, where the errors only come from discretization in numerically solving differential equations and the accuracies in the initial values. Hence our method works even for small degrees of freedom.

The organization of this paper is as follows. In Section 2 we summarize preliminary facts on the exact distribution of the largest root of a Wishart matrix. In particular we state the partial differential equation for ${}_1F_1$ by Muirhead [16]. In Section 3, for expository purposes, we fully describe our holonomic gradient method for dimension two. In Section 4 we derive properties of Pfaffian system for general dimensions. The Pfaffian system is a system of partial differential equations and is called an integrable connection in some literatures. Results of symbolic computations are presented in Section 5 and results of numerical experiments are presented in Section 6. We end the paper with discussion of open problems in Section 7.

2 Preliminaries

Let $\kappa = (k_1, \dots, k_l) \vdash k$ be a partition of a non-negative integer k and define the Pochhammer symbol $(a)_\kappa$ by

$$(a)_\kappa = \prod_{i=1}^{l} \left(a - \frac{i-1}{2} \right)_{k_i}, \quad (a)_{k_i} = \prod_{j=1}^{k_i} (a + j - 1) \quad ((a)_0 = 1).$$

Let $\mathcal{C}_\kappa(Y)$ denote the ("C-normalization" of) zonal polynomial indexed by κ of an $m \times m$ symmetric matrix Y. it is a homogeneous symmetric polynomial of degree k in the characteristic roots $y_1, \dots, y_m$ of Y, satisfying $\sum_{\kappa \vdash k} \mathcal{C}_\kappa(Y) = (\operatorname{tr} Y)^k$. For zonal polynomials in statistics see, e.g., James [8], Muirhead [18], Takemura [30] and Mathai et al. [15]. A hypergeometric function of a matrix argument is defined (Constantine [3]) as

$$ {}_pF_q(a_1, \dots, a_p; c_1, \dots, c_q; Y) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{(a_1)_\kappa \dots (a_p)_\kappa}{(c_1)_\kappa \dots (c_q)_\kappa} \frac{\mathcal{C}_\kappa(Y)}{k!}. \tag{1}$$

In this paper we study holonomic gradient method for ${}_1F_1(a; c; Y)$. Let I_m denote the $m \times m$ identity matrix and let $|X|$ denote the determinant of X. For $\Re a > (m+1)/2$, $\Re(b-a) > (m+1)/2$, ${}_1F_1(a; c; Y)$ has the following integral representation

$$ {}_1F_1(a; c; Y) = \frac{\Gamma_m(b)}{\Gamma_m(a)\Gamma_m(c-a)} \int_{0<X<I_m} \exp(\operatorname{tr} XY) |X|^{a-(m+1)/2} |I_m - X|^{c-a-(m+1)/2} dX, \tag{2}$$

where $0 < X < I_m$ means that X and $I_m - X$ are positive definite, $dX = \prod_{i \le j} dx_{ij}$ is the Lebesgue measure of the upper triangular entries of X, and

$$\Gamma_m(a) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=1}^{m} \Gamma \left(a - \frac{i-1}{2} \right).$$

The hypergeometric function ${}_1F_1$ satisfies the the following Kummer relation (see (2.8) of Herz [6], (51) of James [8]):

$$\exp(-\operatorname{tr} Y) {}_1F_1(a; c; Y) = {}_1F_1(c-a, c; -Y). \tag{3}$$

Note that (2) implies that ${}_1F_1$ is an entire function in Y.

The cumulative distribution function of the largest root ℓ_1 of the $m \times m$ Wishart matrix W with n degrees of freedom and the covariance matrix Σ is written as follows

$$\Pr[\ell_1 < x] = C \exp\left(-\frac{x}{2} \operatorname{tr} \Sigma^{-1} \right) x^{\frac{1}{2}nm} {}_1F_1 \left(\frac{m+1}{2}; \frac{n+m+1}{2}; \frac{x}{2}\Sigma^{-1} \right), \tag{4}$$

where

$$C = \frac{\Gamma_m\left(\frac{m+1}{2}\right)}{2^{\frac{1}{2}nm} (\det \Sigma)^{\frac{1}{2}n} \Gamma_m\left(\frac{n+m+1}{2}\right)}.$$

This follows from the results in Section 9 of Constantine [3] and the Kummer relation (3). See also Sugiyama [27].

The following partial differential equations for ${}_1F_1(a;b;Y)$ were derived by Muirhead [16].

Theorem 1 (Theorem 5.1 of Muirhead [16], Theorem 7.5.6 of Muirhead [18])**.** *The hypergeometric function $F = {}_1F_1(a;c;Y)$ of a matrix argument $Y = \mathrm{diag}(y_1,\ldots,y_m)$ is the unique solution of the following set of m partial differential equations*

$$\left[y_i\,\partial_i^2 + \left\{ c - \frac{m-1}{2} - y_i + \frac{1}{2}\sum_{j=1,j\neq i}^{m} \frac{y_i}{y_i - y_j} \right\} \partial_i - \frac{1}{2}\sum_{j=1,j\neq i}^{m} \frac{y_j}{y_i - y_j}\partial_j - a \right] F = 0, \qquad (5)$$

$$(i = 1,\ldots,m),$$

subject to the conditions that F is symmetric in $y_1,\ldots,y_m$ and F is analytic at $Y = 0$, $F(0) = 1$.

The partial differential equation (5) has singularities along $y_i = 0$ and $y_j = y_i$, $j \neq i$. However since F is an entire function, F is determined by the partial differential equations on the open region $\mathcal{X} = \{y \in \mathbb{C}^m \mid \prod_{i=1}^m y_i \prod_{i\neq j}(y_i - y_j) \neq 0\}$. In this paper we call $\mathcal{X}$ the non-diagonal region. Using

$$\frac{y_i}{y_i - y_j} = 1 + \frac{y_j}{y_i - y_j}$$

we can rewrite (5) as $g_i F = 0$, $i = 1,\ldots,m$, where

$$g_i = y_i \partial_i^2 + (c - y_i)\partial_i + \frac{1}{2}\sum_{j=1,j\neq i}^{m} \frac{y_j}{y_i - y_j}(\partial_i - \partial_j) - a \qquad (6)$$

is a differential operator annihilating F. In our holonomic gradient method we make a direct use of the partial differential equations for numerical evaluation of ${}_1F_1$.

3 Holonomic gradient method for dimension two

In this section we illustrate the holonomic gradient method for the case of $m = 2$. Although our purpose is to implement an algorithm of our method for a larger dimension, for clarity it is best to do "by hand" calculation for the case of $m = 2$. As in the previous section we simply write $F(Y) = {}_1F_1(a;c;Y)$.

In Nakayama et al. [19] the holonomic gradient method was used to obtain the maximum likelihood estimate. The reciprocal of the likelihood function was minimized and the method was called the holonomic gradient *descent*. For the application of this paper we simply use the holonomic gradient method for evaluating F. Hence we omit the term "descent". Also, for minimization, at each step of the iteration, a direction for increments was chosen to decrease the value of the function. In our application, starting from the

origin $Y = 0$, we can choose arbitrary path to the target value Y where we want to evaluate $F(Y)$.

Another minor difference of the expository explanation in this section from Nakayama et al. [19] and Sei et al. [24] is that we use the simple forward Euler method (e.g., Section 3.1 of Ascher and Petzold [1]) for updating partial derivatives of F. In Nakayama et al. [19], once an updating direction is chosen at each step of the iteration, the 4-th order Runge-Kutta method was used. The simple Euler method is used *only for the purpose of exposition*. It is easier to explain the basic idea of the holonomic gradient method with the simple Euler method. In our actual implementation in Section 6 we use the Runge-Kutta method for numerically solving the differential equation.

We will reduce our problem to a traditional problem of numerical analysis of an ordinary differential equation (ODE). For the reduction we utilize the notion of holonomic differential equations and the gradients of their solutions. It is why we call our method holonomic gradient method.

In the following we discuss the case of $y_1 \neq y_2$ and $y_1 = y_2$ separately.

3.1 Holonomic gradient method for non-diagonal region

In this subsection we assume $y_1 \neq y_2$. Two partial differential equations in (6) are written as

$$\Big[y_1\partial_1^2 + (c - y_1)\partial_1 + \frac{1}{2}\frac{y_2}{y_1 - y_2}(\partial_1 - \partial_2) - a\Big]F = 0, \tag{7}$$

$$\Big[y_2\partial_2^2 + (c - y_2)\partial_2 + \frac{1}{2}\frac{y_1}{y_2 - y_1}(\partial_2 - \partial_1) - a\Big]F = 0. \tag{8}$$

Suppose that we want to evaluate a higher derivative $\partial_1^{n_1}\partial_2^{n_2}F = \partial_2^{n_2}\partial_1^{n_1}F$ of F. Let $n_2 \geq 2$. Then by (8)

$$\partial_1^{n_1}\partial_2^{n_2}F = \partial_1^{n_1}\partial_2^{n_2-2}\big(- \frac{c}{y_2}\partial_2 + \partial_2 - \frac{1}{2}\frac{y_1}{y_2(y_2 - y_1)}(\partial_2 - \partial_1) + \frac{a}{y_2}\big)F. \tag{9}$$

Noting

$$\partial_2\frac{1}{y_2} = -\frac{1}{y_2^2}, \quad \partial_2\frac{y_1}{y_2(y_2 - y_1)} = -\frac{y_1(2y_2 - y_1)}{y_2^2(y_2 - y_1)^2},$$

for $n_2 > 2$, the right-hand side of (9) is further written as

$$\begin{aligned}\partial_1^{n_1}\partial_2^{n_2-3}\Big(&\frac{c}{y_2^2}\partial_2 - \frac{c - y_2}{y_2}\partial_2^2 + \frac{1}{2}\frac{y_1(2y_2 - y_1)}{y_2^2(y_2 - y_1)^2}(\partial_2 - \partial_1)\\ &- \frac{1}{2}\frac{y_1}{y_2(y_2 - y_1)}(\partial_2^2 - \partial_1\partial_2) - \frac{a}{y_2^2} + \frac{a}{y_2}\partial_2\Big)F.\end{aligned} \tag{10}$$

Although the result is somewhat complicated, the important fact is that the total degree of differentiation $n_1 + n_2$ on the left-hand side of (9) is decreased by one to $n_1 + n_2 - 1$ in (10). As long as the degree of ∂_1 or ∂_2 is more than one, then we can recursively apply

(7) or (8) to decrease the total degree of differentiation. It follows that for each n_1, n_2, there exist rational functions $h_{00}^{(n_1,n_2)}, h_{10}^{(n_1,n_2)}, h_{01}^{(n_1,n_2)}, h_{11}^{(n_1,n_2)}$ in (y_1, y_2) such that

$$\partial_1^{n_1}\partial_2^{n_2}F = h_{00}^{(n_1,n_2)}F + h_{10}^{(n_1,n_2)}\partial_1 F + h_{01}^{(n_1,n_2)}\partial_2 F + h_{11}^{(n_1,n_2)}\partial_1\partial_2 F. \tag{11}$$

In this notation (7) is written as

$$\begin{aligned}\partial_1^2 F &= \frac{a}{y_1}F - (\frac{c-y_1}{y_1} + \frac{1}{2}\frac{y_2}{y_1(y_1-y_2)})\partial_1 F + \frac{1}{2}\frac{y_2}{y_1(y_1-y_2)}\partial_2 F \\ &= h_{00}^{(2,0)}F + h_{10}^{(2,0)}\partial_1 F + h_{01}^{(2,0)}\partial_2 F \qquad (h_{11}^{(2,0)} \equiv 0).\end{aligned} \tag{12}$$

For a general dimension, (11) corresponds to the reduction by a Gröbner basis as discussed in Section 4.

For us the important case is $n_1 = 1, n_2 = 2$. Since

$$\partial_1 \frac{y_1}{y_2(y_2-y_1)} = \partial_1\big(\frac{1}{y_2-y_1} - \frac{1}{y_2}\big) = \frac{1}{(y_2-y_1)^2}$$

we have

$$\begin{aligned}\partial_1\partial_2^2 F &= \partial_1\big(-\frac{c-y_2}{y_2}\partial_2 - \frac{1}{2}\frac{y_1}{y_2(y_2-y_1)}(\partial_2-\partial_1) + \frac{a}{y_2}\big)F \\ &= \big(-\frac{c-y_2}{y_2}\partial_1\partial_2 - \frac{1}{2}\frac{1}{(y_2-y_1)^2}(\partial_2-\partial_1) - \frac{1}{2}\frac{y_1}{y_2(y_2-y_1)}(\partial_1\partial_2 - \partial_1^2) + \frac{a}{y_2}\partial_1\big)F.\end{aligned}$$

There is a term $y_1\partial_1^2$ on the right-hand side, into which we further substitute (7). Then (11) for $\partial_1\partial_2^2 F$ is written as

$$\begin{aligned}\partial_1\partial_2^2 F &= \Big(-\frac{c-y_2}{y_2}\partial_1\partial_2 - \frac{1}{2}\frac{1}{(y_2-y_1)^2}(\partial_2-\partial_1) - \frac{1}{2}\frac{y_1}{y_2(y_2-y_1)}\partial_1\partial_2 + \frac{a}{y_2}\partial_1 \\ &\qquad - \frac{1}{2y_2(y_2-y_1)}\big((c-y_1)\partial_1 + \frac{1}{2}\frac{y_2}{y_1-y_2}(\partial_1-\partial_2) - a\big)\Big)F \\ &= \frac{a}{2y_2(y_2-y_1)}F + \Big(\frac{3}{4}\frac{1}{(y_2-y_1)^2} + \frac{a}{y_2} - \frac{c-y_1}{2y_2(y_2-y_1)}\Big)\partial_1 F \\ &\qquad - \frac{3}{4}\frac{1}{(y_2-y_1)^2}\partial_2 F - \Big(\frac{c-y_2}{y_2} + \frac{1}{2}\frac{y_1}{y_2(y_2-y_1)}\Big)\partial_1\partial_2 F \\ &= h_{00}^{(1,2)}F + h_{10}^{(1,2)}\partial_1 F + h_{01}^{(1,2)}\partial_2 F + h_{11}^{(1,2)}\partial_1\partial_2 F.\end{aligned} \tag{13}$$

Since F is a symmetric function in y_1 and y_2, $\partial_1^2\partial_2 F$ is obtained by permuting y_1 and y_2.

Let

$$\vec{F} = \begin{pmatrix} F \\ \partial_1 F \\ \partial_2 F \\ \partial_1\partial_2 F\end{pmatrix}$$

denote the vector consisting of F and its square-free mixed derivatives. Differentiate the components of $\vec{F}$ by y_1 and denote $\partial_1\vec{F} = (\partial_1 F, \partial_1^2 F, \partial_1\partial_2 F, \partial_1^2\partial_2 F)^t$. Similarly define $\partial_2\vec{F}$. Then by (12) and (13), $\partial_i\vec{F}$, $i = 1, 2$, are written as $\partial_i\vec{F} = P_i(Y)\vec{F}$, where P_1 and P_2 are the following 4×4 matrices with rational function entries

$$P_1(Y) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ h_{00}^{(2,0)} & h_{10}^{(2,0)} & h_{01}^{(2,0)} & 0 \\ 0 & 0 & 0 & 1 \\ h_{00}^{(2,1)} & h_{10}^{(2,1)} & h_{01}^{(2,1)} & h_{11}^{(2,1)} \end{pmatrix}, \quad P_2(Y) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ h_{00}^{(0,2)} & h_{10}^{(0,2)} & h_{01}^{(0,2)} & 0 \\ h_{00}^{(1,2)} & h_{10}^{(1,2)} & h_{01}^{(1,2)} & h_{11}^{(1,2)} \end{pmatrix}.$$

The matrices P_1, P_2 are called coefficient matrices of a *Pfaffian system* (an integrable connection) for F (Nakayama et al. [19]). Note that P_2 is obtained from P_1 by permutation of y_1 and y_2. If we know the values of the components of $\vec{F}$ at $Y = (y_1, y_2)$, $y_1 \neq y_2$, then values at a nearby point $Y + \Delta Y = (y_1 + \Delta y_1, y_2 + \Delta y_2)$ can be approximated by the simple Euler method (i.e. linear approximation) as

$$\begin{aligned} \vec{F}(Y + \Delta Y) &\doteq \vec{F}(Y) + \Delta y_1 \partial_1 \vec{F}(Y) + \Delta y_2 \partial_2 \vec{F}(Y) \\ &= \vec{F}(Y) + \Delta y_1 P_1(Y)\vec{F}(Y) + \Delta y_2 P_2(Y)\vec{F}(Y). \end{aligned} \tag{14}$$

Now suppose that we want to evaluate $F(y_1, y_2)$ at a particular point (y_1, y_2) with $y_1 \neq y_2$. If we know $\vec{F}(Y_0)$ at some point $Y_0 = (y_1^{(0)}, y_2^{(0)})$, $y_1^{(0)} \neq y_2^{(0)}$, close to the origin, then we can choose an appropriate sequence of points $Y^{(l)} = (y_1^{(l)}, y_2^{(l)})$, $l = 0, \ldots, L$, such that $(y_1, y_2) = (y_1^{(L)}, y_2^{(L)})$. Along the sequence we can use (14) to update $\vec{F}(Y^{(l)})$ and finally the first element of $\vec{F}(Y^{(L)})$ gives $F(y_1, y_2)$.

Therefore it remains to consider how to obtain the initial values. Close to the origin we can use the definition (1) of ${}_1F_1$. If Y is very close to zero, then we only need zonal polynomials of low orders, whose explicit forms are known. Zonal polynomials up to the third order are as follows; $\mathcal{C}_{(1)}(Y) = \mathcal{M}_{(1)}(Y)$,

$$\begin{pmatrix} \mathcal{C}_{(2)}(Y) \\ \mathcal{C}_{(1,1)}(Y) \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{3} \\ 0 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} \mathcal{M}_{(2)}(Y) \\ \mathcal{M}_{(1,1)}(Y) \end{pmatrix}, \quad \begin{pmatrix} \mathcal{C}_{(3)}(Y) \\ \mathcal{C}_{(2,1)}(Y) \\ \mathcal{C}_{(1,1,1)}(Y) \end{pmatrix} = \begin{pmatrix} 1 & \frac{3}{5} & \frac{2}{5} \\ 0 & \frac{12}{5} & \frac{18}{5} \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mathcal{M}_{(3)}(Y) \\ \mathcal{M}_{(2,1)}(Y) \\ \mathcal{M}_{(1,1,1)}(Y) \end{pmatrix}, \tag{15}$$

where $\mathcal{M}_\kappa(Y)$ is the monomial symmetric polynomial associated with a partition κ. Since $F(y_1, y_2)$ can be expanded as

$$\begin{aligned} F(y_1, y_2) &= 1 + \frac{(a)_{(1)}}{(c)_{(1)}}\mathcal{C}_{(1)}(Y) + \frac{1}{2!}\left(\frac{(a)_{(2)}}{(c)_{(2)}}\mathcal{C}_{(2)}(Y) + \frac{(a)_{(1,1)}}{(c)_{(1,1)}}\mathcal{C}_{(1,1)}(Y)\right) + \cdots \\ &= 1 + \frac{(a)_{(1)}}{(c)_{(1)}}\mathcal{M}_{(1)}(Y) + \frac{(a)_{(2)}}{2(c)_{(2)}}\mathcal{M}_{(2)}(Y) + \left(\frac{(a)_{(2)}}{3(c)_{(2)}} + \frac{2(a)_{(1,1)}}{3(c)_{(1,1)}}\right)\mathcal{M}_{(1,1)}(Y) + \cdots, \end{aligned} \tag{16}$$

for an example, $\partial_1\partial_2 F(0,0)$ is obtained as

$$\partial_1\partial_2 F(0,0) = \frac{(a)_2}{3(c)_2} + \frac{2a(a-\frac{1}{2})}{3c(c-\frac{1}{2})}.$$

In a similar manner, we have

$$\partial_1 F(0,0) = \partial_2 F(0,0) = \frac{a}{c},\ \partial_1^2 F(0,0) = \partial_2^2 F(0,0) = \frac{(a)_2}{(c)_2},$$
$$\partial_1^2\partial_2 F(0,0) = \partial_2^2\partial_1 F(0,0) = \frac{(a)_3}{5(c)_3} + \frac{4(a)_2(a-\frac{1}{2})}{5(c)_2(c-\frac{1}{2})}. \tag{17}$$

These formulae can be obtained by a symbolic mathematics software, such as the routines for Jack polynomials in `sage` mathematics software system (Stein et al. [26]).

In order to obtain the initial value $\vec{F}(Y_0)$ at $Y_0 = (y_1^{(0)}, y_2^{(0)})$ close to the origin, we can use the approximation

$$\vec{F}(y_1^{(0)}, y_2^{(0)}) \doteq \vec{F}(0,0) + y_1^{(0)}\partial_1\vec{F}(0,0) + y_2^{(0)}\partial_2\vec{F}(0,0). \tag{18}$$

We code the above procedure using deSolve package in the data analysis system R. We show a simple source program in Appendix B. In addition, since the zonal polynomials are easy to evaluate for $m=2$, we also evaluate the series expansion of ${}_1F_1$ up to $k=150$. As an example, we compute percentage points by two methods for the case of $n=3$, $\Sigma = \mathrm{diag}(1/2, 1/4)$. The following percentage points for ℓ_1 agree in two methods to 6 digits.

50%	90%	95%	99%
1.63785	3.54999	4.31600	6.05836

3.2 Holonomic gradient method for the diagonal line

In the previous subsection we assumed $y_1 \neq y_2$ to avoid singularity of the differential equations. However ${}_1F_1$ itself does not have singularities. Hence we should be able to derive some differential equation even for $y = y_1 = y_2$.

In (7) and (8) we can perform the limiting operation $y_1 \to y_2 = y$ using the l'Hôpital rule. Since F is a symmetric function, at (y,y) we have

$$\partial_1 F(y,y) = \partial_2 F(y,y).$$

Also $\partial_1^2 F(y,y) = \partial_2^2 F(y,y)$. Hence by the l'Hôpital rule, in (7) we have

$$\lim_{y_1\to y_2=y} \frac{\partial_1 F - \partial_2 F}{y_1 - y_2} = \partial_2^2 F - \partial_1\partial_2 F = \partial_1^2 F - \partial_1\partial_2 F.$$

Then (7) for at (y,y) is written as

$$0 = \left[y\partial_1^2 + (c-y)\partial_1 + \frac{y}{2}(\partial_1^2 - \partial_1\partial_2) - a\right]F = \left[\frac{3}{2}y\partial_1^2 + (c-y)\partial_1 - \frac{y}{2}\partial_1\partial_2 - a\right]F. \tag{19}$$

Based on this we derive an ODE for $f(y) = F(y, y)$. Firstly,

$$f'(y) = 2\partial_1 F \quad \text{or} \quad \partial_1 F = f'(y)/2.$$

Secondly,

$$f''(y) = 2\partial_1^2 F + 2\partial_1\partial_2 F.$$

From (19)

$$\frac{3}{2} y\partial_1^2 F = \frac{1}{2} y\partial_1\partial_2 F - (c - y)\partial_1 F + aF,$$

and

$$\begin{aligned}
\frac{3}{4} yf''(y) &= \frac{1}{2} y\partial_1\partial_2 F - (c - y)\partial_1 F + aF + \frac{3}{2} y\partial_1\partial_2 F \\
&= 2y\partial_1\partial_2 F - (c - y)\partial_1 F + aF \\
&= 2y\partial_1\partial_2 F - \frac{c - y}{2} f'(y) + af(y)
\end{aligned}$$

or

$$\partial_1\partial_2 F(y, y) = \frac{3}{8} f''(y) + \frac{c - y}{4y} f'(y) - \frac{a}{2y} f(y). \tag{20}$$

Thirdly,

$$f'''(y) = 2\partial_1^3 F + 6\partial_1^2\partial_2 F. \tag{21}$$

In order to get another relation for $\partial_1^3 F$ and $\partial_1^2\partial_2 F$, we differentiate (7) by y_2. Then by

$$\partial_2 \frac{y_2}{y_1 - y_2} = \partial_2\left(\frac{y_1}{y_1 - y_2} - 1\right) = \frac{y_1}{(y_1 - y_2)^2}, \tag{22}$$

we obtain the following differential operator annihilating F:

$$y_1\partial_1^2\partial_2 + (c - y_1)\partial_1\partial_2 + \frac{1}{2}\frac{y_1}{(y_1 - y_2)^2}(\partial_1 - \partial_2) + \frac{1}{2}\frac{y_2}{y_1 - y_2}(\partial_1\partial_2 - \partial_2^2) - a\partial_2. \tag{23}$$

Noting $y_2/(y_1 - y_2) = y_1/(y_1 - y_2) - 1$ this can be further written as

$$\begin{aligned}
&y_1\partial_1^2\partial_2 + (c - y_1)\partial_1\partial_2 + \frac{y_1}{2}\frac{(\partial_1 - \partial_2) + (y_1 - y_2)(\partial_1\partial_2 - \partial_2^2)}{(y_1 - y_2)^2} - \frac{1}{2}(\partial_1\partial_2 - \partial_2^2) - a\partial_2 \\
&= y_1\partial_1^2\partial_2 + (c - 1 - y_1)\partial_1\partial_2 + \frac{y_1}{2}\frac{(\partial_1 - \partial_2) + (y_1 - y_2)(\partial_1\partial_2 - \partial_2^2)}{(y_1 - y_2)^2} + \frac{1}{2}(\partial_1\partial_2 + \partial_2^2) - a\partial_2.
\end{aligned}$$

We now apply the l'Hôpital rule to

$$\frac{(\partial_1 - \partial_2) + (y_1 - y_2)(\partial_1\partial_2 - \partial_2^2)}{(y_1 - y_2)^2}.$$

We again let $y_1 \to y_2 = y$. The second derivative of the denominator with respect to y_1 gives 2. Now

$$
\begin{aligned}
&\partial_1^2\big((\partial_1 - \partial_2) + (y_1 - y_2)(\partial_1\partial_2 - \partial_2^2)\big) \\
&\quad = \partial_1\big((\partial_1^2 - \partial_1\partial_2) + (\partial_1\partial_2 - \partial_2^2) + (y_1 - y_2)(\partial_1^2\partial_2 - \partial_1\partial_2^2)\big) \\
&\quad = \partial_1\big((\partial_1^2 - \partial_2^2) + (y_1 - y_2)(\partial_1^2\partial_2 - \partial_1\partial_2^2)\big) \\
&\quad = (\partial_1^3 - \partial_1\partial_2^2) + (\partial_1^2\partial_2 - \partial_1\partial_2^2) + (y_1 - y_2)(\partial_1^3\partial_2 - \partial_1^2\partial_2^2).
\end{aligned}
$$

Evaluating the right-hand side at $y = y_1 = y_2$ and noting that $\partial_1^2\partial_2 F = \partial_1\partial_2^2 F$ at (y, y), we just have $\partial_1^3 - \partial_1^2\partial_2$. Hence (23) at (y, y) reduces to

$$
\begin{aligned}
&y\partial_1^2\partial_2 + (c - 1 - y)\partial_1\partial_2 + \frac{y}{4}(\partial_1^3 - \partial_1^2\partial_2) + \frac{1}{4}(2\partial_1^2 + 2\partial_1\partial_2) - a\partial_1 \\
&\quad = \frac{y}{8}(2\partial_1^3 + 6\partial_1^2\partial_2) + (c - 1 - y)\partial_1\partial_2 + \frac{1}{4}(2\partial_1^2 + 2\partial_1\partial_2) - a\partial_1,
\end{aligned}
$$

where we used $\partial_1 F = \partial_2 F$ at (y, y). Comparing the right-hand side with (21) and by (20) we obtain

$$
\frac{y}{8}f'''(y) + (c - 1 - y)\Big(\frac{3}{8}f''(y) + \frac{c - y}{4y}f'(y) - \frac{a}{2y}f(y)\Big) + \frac{1}{4}f''(y) - \frac{a}{2}f'(y) = 0. \tag{24}
$$

This equation can be written as

$$
f'''(y) = h_2(y)f''(y) + h_1(y)f'(y) + h_0(y)f(y),
$$

where

$$
h_2(y) = -\frac{3(c - 1 - y)}{y} - \frac{2}{y},\ h_1(y) = \frac{4a}{y} - \frac{2(c - y)(c - 1 - y)}{y^2},\ h_0(y) = \frac{4a(c - 1 - y)}{y^2}
$$

are rational functions in y. The coefficient matrix for the Pfaffian system for a one-dimensional ODE is simply the companion matrix

$$
P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ h_0(y) & h_1(y) & h_2(y) \end{pmatrix}.
$$

Note that the values of f, f', f'' and f''' at the origin are given by

$$
\begin{aligned}
f(0) &= F(0,0) = 1, \quad f'(0) = 2\partial_1 F(0,0) = \frac{2a}{c}, \\
f''(0) &= 2\partial_1^2 F(0,0) + 2\partial_1\partial_2 F(0,0) = \frac{8(a)_2}{3(c)_2} + \frac{4a(a - \frac{1}{2})}{3c(c - \frac{1}{2})}, \\
f'''(0) &= 2\partial_1^3 F(0,0) + 6\partial_1^2\partial_2 F(0,0) = 2\frac{(a)_3}{(c)_3} + 6\Big(\frac{(a)_3}{5(c)_3} + \frac{4(a)_2(a - \frac{1}{2})}{5(c)_2(c - \frac{1}{2})}\Big).
\end{aligned}
$$

As seen above, the computation using the l'Hôpital rule is already tedious for $m = 2$. Actually the computation can be automated by the restriction algorithm for holonomic ideals. This will be explained in Section 5.2.

4 Properties of the Pfaffian system (integrable connection) for a general dimension

We now consider our problem for a general dimension. We fully utilize Gröbner basis theory for the ring of differential operators. In this section we only consider the non-diagonal region $\mathcal{X}$. Let $K = \mathbb{C}(y_1, \ldots, y_m)$ be the field of rational functions in $y_1, \ldots, y_m$ with complex coefficients. Further let

$$R = K\langle \partial_1, \ldots, \partial_m \rangle = \mathbb{C}(y_1, \ldots, y_m)\langle \partial_1 \ldots, \partial_m \rangle$$

be the ring of differential operators with rational function coefficients (see Appendix of Nakayama et al. [19]). Let I denote the left ideal of R generated by $g_1, \ldots, g_m$:

$$I = \langle g_1, \ldots, g_m \rangle, \tag{25}$$

where g_i is given in (6).

We now prove the following lemma concerning the commutators of $g_1, \ldots, g_m$.

Lemma 1. *For* $1 \le i \ne j \le m$,

$$[g_i, g_j] = -\frac{1}{2} \frac{y_i + y_j}{(y_i - y_j)^2}(g_i - g_j). \tag{26}$$

A similar result for ${}_2F_1$ is given in Lemma 9.9 of Ibukiyama et al. [7]. Although they claim that their Lemma 9.9 follows from a straightforward computation, in fact the computation for checking (26) is tedious even for $m = 2$. However for $m = 2$, (26) can be verified by some software systems (e.g., RisaAsir developing team [22]), which can handle rings of differential operators. The following program in `Risa/Asir`

```
import("names.rr"); import("yang.rr");
yang.define_ring(["partial",[y1,y2]]);
G1=y1*dy1^2+(c-y1)*dy1+(1/2)*(y2/(y1-y2))*(dy1-dy2)-a;
G2=base_replace(G1,[[y1,y2],[y2,y1],[dy1,dy2],[dy2,dy1]]);
G=yang.mul(G1,G2)-yang.mul(G2,G1)+(1/2)*(y1+y2)/(y1-y2)^2*(G1-G2);
printf("G=%a\n",G);
```

outputs the result `G=0`. Therefore in the following proof, assuming that (26) holds for $m = 2$, we show that it holds for $m > 2$.

Proof. By symmetry we only need to prove the case $i = 1, j = 2$. Define $\tilde{g}_1, \tilde{g}_2$

$$\tilde{g}_1 = y_1\partial_1^2 + (c - y_1)\partial_1 + \frac{1}{2}\frac{y_2}{y_1 - y_2}(\partial_1 - \partial_2) - a,$$
$$\tilde{g}_2 = y_2\partial_2^2 + (c - y_2)\partial_2 + \frac{1}{2}\frac{y_1}{y_2 - y_1}(\partial_2 - \partial_1) - a.$$

Then

$$g_i = \tilde{g}_i + h_i, \quad h_i = \frac{1}{2}\sum_{k=3}^{m} \frac{y_k}{y_i - y_k}(\partial_i - \partial_k), \quad i = 1, 2.$$

We already know

$$[\tilde{g}_1, \tilde{g}_2] = -\frac{1}{2}\frac{y_1 + y_2}{(y_1 - y_2)^2}(\tilde{g}_1 - \tilde{g}_2).$$

Then

$$[g_1, g_2] = [\tilde{g}_1 + h_1, \tilde{g}_2 + h_2] = [\tilde{g}_1, \tilde{g}_2] + [h_1, \tilde{g}_2] + [\tilde{g}_1, h_2] + [h_1, h_2].$$

Therefore it suffices to show

$$[h_1, \tilde{g}_2] + [\tilde{g}_1, h_2] + [h_1, h_2] = -\frac{1}{2}\frac{y_1 + y_2}{(y_1 - y_2)^2}(h_1 - h_2).$$

In considering commutators, we only need to look at terms, where a differential operator actually differentiate rational functions in $y_1, \ldots, y_m$. For example consider $h_1\tilde{g}_2$ in $[h_1, \tilde{g}_2]$. In $h_1\tilde{g}_2$ the only relevant term is ∂_1 in h_1 differentiating $y_1/(y_1 - y_2)$ in $\tilde{g}_2$. Noting

$$\partial_1 \frac{y_1}{y_2 - y_1} = \partial_1 \Big(\frac{y_2}{y_2 - y_1} - 1\Big) = \frac{y_2}{(y_2 - y_1)^2},$$

in $h_1\tilde{g}_2$ the relevant terms are

$$\frac{1}{4}\frac{y_2}{(y_2 - y_1)^2}\sum_{k=3}^{m}\frac{y_k}{y_1 - y_k}(\partial_2 - \partial_1) = \frac{1}{4}\frac{y_2}{(y_2 - y_1)^2}\sum_{k=3}^{m}\frac{y_k}{y_1 - y_k}\big((\partial_2 - \partial_k) - (\partial_1 - \partial_k)\big).$$

In $\tilde{g}_2 h_1$ we need to look at ∂_1 in $\tilde{g}_2$ differentiating $y_k/(y_1 - y_k)$. Hence we have

$$\frac{1}{4}\frac{y_1}{y_2 - y_1}\sum_{k=3}^{m}\frac{y_k}{(y_1 - y_k)^2}(\partial_1 - \partial_k).$$

Similarly in $[\tilde{g}_1, h_2] = -[h_2, \tilde{g}_1]$ the relevant terms are

$$-\frac{1}{4}\frac{y_1}{(y_1 - y_2)^2}\sum_{k=3}^{m}\frac{y_k}{y_2 - y_k}\big((\partial_1 - \partial_k) - (\partial_2 - \partial_k)\big) + \frac{1}{4}\frac{y_2}{y_1 - y_2}\sum_{k=3}^{m}\frac{y_k}{(y_2 - y_k)^2}(\partial_2 - \partial_k).$$

Finally in $[h_1, h_2]$ we look at ∂_k differentiating $y_k/(y_i - y_k)$. Then the relevant terms are

$$-\frac{1}{4}\sum_{k=3}^{m}\frac{y_k}{y_1 - y_k}\frac{y_2}{(y_2 - y_k)^2}(\partial_2 - \partial_k) + \frac{1}{4}\sum_{k=3}^{m}\frac{y_k}{y_2 - y_k}\frac{y_1}{(y_1 - y_k)^2}(\partial_1 - \partial_k).$$

Then the coefficient for $-(\partial_1 - \partial_k)/4$ is

$$\begin{aligned} &\frac{y_2}{(y_2 - y_1)^2}\frac{y_k}{y_1 - y_k} + \frac{y_1}{y_2 - y_1}\frac{y_k}{(y_1 - y_k)^2} + \frac{y_1}{(y_1 - y_2)^2}\frac{y_k}{y_2 - y_k} - \frac{y_k}{y_2 - y_k}\frac{y_1}{(y_1 - y_k)^2} \\ &\quad = \frac{y_1 + y_2}{(y_2 - y_1)^2}\frac{y_k}{y_1 - y_k}, \end{aligned}$$

which coincides with the coefficient of $-(\partial_1 - \partial_k)/4$ in

$$-\frac{1}{2}\frac{y_1 + y_2}{(y_1 - y_2)^2}h_1.$$

Similarly the coefficients of $(\partial_2 - \partial_k)$ coincide on both sides. $\square$

We now consider the graded lexicographic term order $\succ$. The initial term of g_i (without the coefficient y_i) is given as

$$\mathrm{in}_{\succ} g_i = \partial_i^2.$$

We now prove the following theorem.

Theorem 2. *For the term order $\succ$, $\{g_1, \ldots, g_m\}$ is a Gröbner basis of I in R and the initial ideal is given by $\langle \partial_1^2, \ldots, \partial_m^2 \rangle$. I is zero-dimensional and the set of standard monomials is given by the set of square-free mixed derivatives*

$$\{\partial_{i_1}\partial_{i_2} \ldots \partial_{i_k} \mid \ 1 \le i_1 < \cdots < i_k \le m, \ k \le m\},$$

which has the cardinality 2^m.

Proof. By Lemma 1 and the Buchberger's criterion for the ring R (cf. Theorem 1.1.10 of Saito et al. [23]), $g_i, i = 1, \ldots, m$, form a Gröbner basis and the initial ideal is given by $\langle \partial_1^2, \ldots, \partial_m^2 \rangle$. Let $J = \langle \partial_1, \ldots, \partial_m \rangle$. Then $J^{m+1} \subset \langle \partial_1^2, \ldots, \partial_m^2 \rangle$. Hence I is a zero-dimensional ideal. Furthermore this shows that the set of standard monomials is given by the set of square-free mixed derivatives. $\square$

It follows from Theorem 2 that there exists a Pfaffian system and $2^m \times 2^m$ matrices (as $P_i(Y)$ for $m = 2$ in the expository section 3) are obtained by the normal form algorithm in the ring of differential operators R. The matrices are used to numerically solve the associated ODE. However, the derivation of the matrices on computer is heavy and the obtained matrices are not in a relevant form for an efficient numerical evaluation. Then, we do it *by hand* in the sequel.

Consider a higher order derivative $\partial_1^{n_1} \ldots \partial_m^{n_m} F$ of $F = {}_1F_1(a; c; y_1, \ldots, y_m)$. If total degree of differentiation $n = n_1 + \cdots + n_m$ is greater than or equal to $m + 1$, then for some i we have $n_i \ge 2$. Then as in the previous section we can use $g_i F = 0$ to decrease the total degree of differentiation. Therefore as in (11), for each $n_1, \ldots, n_m$, there exist 2^m rational functions $h_{i_1,\ldots,i_m}^{(n_1,\ldots,n_m)}$, $i_j = 0, 1$, $j = 1, \ldots, m$, such that

$$\partial_1^{n_1} \ldots \partial_m^{n_m} F = \sum_{i_1=0}^{1} \cdots \sum_{i_m=0}^{1} h_{i_1,\ldots,i_m}^{(n_1,\ldots,n_m)} \partial_1^{i_1} \ldots \partial_m^{i_m} F. \tag{27}$$

In the holonomic gradient method, as in the case of $m = 2$ in (14), we only need $h_{i_1,\ldots,i_m}^{(n_1,\ldots,n_m)}$ where $0 \le n_1, \ldots, n_m \le 2$ and at most one of $n_1, \ldots, n_m$ is two, such as $h_{i_1,\ldots,i_m}^{(2,1,\ldots,1,0,\ldots,0)}$. Define a 2^m-dimensional vector of square-free mixed derivatives of F by $\vec{F} = (F, \partial_1 F, \partial_2 F, \partial_1\partial_2 F, \ldots, \partial_1 \ldots \partial_m F)^t$. In $\vec{F}$ the elements are lexicographically ordered, for convenience in programming. $\partial_i \vec{F}$ is written as

$$\partial_i \vec{F} = P_i(y)\vec{F}, \qquad i = 1, \ldots, m,$$

where $P_i(y)$, $i = 1, \ldots, m$, in the Pfaffian system are $2^m \times 2^m$ matrices consisting of $h_{i_1,\ldots,i_m}^{(n_1,\ldots,n_m)}$'s.

We now study the form of $h^{(n_1,\dots,n_m)}_{i_1,\dots,i_m}$, where $0 \le n_1,\dots,n_m \le 2$ and at most one of $n_1,\dots,n_m$ is two. Denote $[m] = \{1,\dots,m\}$. For a subset $J \subset [m]$ denote

$$\partial_J = \prod_{j\in J} \partial_j.$$

Choose $i \in [m]$ and $J \subset [m]$ such that $i \notin J$. Write $I = J \cup \{i\}$. $\partial_J g_i F = \partial_J 0 = 0$, where g_i is in (6). Since $i \notin J$, we can write $\partial_J g_i$ as

$$y_i\partial_i^2\partial_J + (c-y_i)\partial_I + \frac{1}{2}\sum_{k\neq i}\partial_J\big(\frac{y_k}{y_i-y_k}(\partial_i-\partial_k)\big) - a\partial_J.$$

For $k \notin J$

$$\partial_J\big(\frac{y_k}{y_i-y_k}(\partial_i-\partial_k)\big) = \frac{y_k}{y_i-y_k}(\partial_I - \partial_{J\cup\{k\}}).$$

On the other hand for $k \in J$, by (22)

$$\partial_J\big(\frac{y_k}{y_i-y_k}(\partial_i-\partial_k)\big) = \frac{y_k}{y_i-y_k}(\partial_I - \partial_J\partial_k) + \frac{y_i}{(y_i-y_k)^2}(\partial_{\{i\}\cup J\setminus\{k\}} - \partial_J).$$

Here $\partial_J\partial_k$ is not square-free and in fact

$$\partial_J\partial_k = \partial_k^2\partial_{J\setminus\{k\}},$$

which causes recursive application of (6). In $\partial_J g_i$ we now separate square-free terms and define

$$\begin{aligned} r(i,J;y) = -\Big[&(c-y_i)\partial_I - a\partial_J + \frac{1}{2}\sum_{k\notin I}\frac{y_k}{y_i-y_k}(\partial_I - \partial_J\partial_k) \\ &+ \frac{1}{2}\sum_{k\in J}\frac{y_k}{y_i-y_k}\partial_I + \frac{1}{2}\sum_{k\in J}\frac{y_i}{(y_i-y_k)^2}(\partial_i\partial_{J\setminus\{k\}} - \partial_J)\Big], \end{aligned}$$

where for $J = \emptyset$, reflecting the original g_i, we define

$$r(i,\emptyset;y) = -\Big[(c-y_i)\partial_i - a + \frac{1}{2}\sum_{k\neq i}\frac{y_k}{y_i-y_k}(\partial_i - \partial_k)\Big].$$

Then $\partial_i^2\partial_J F$ is expanded as

$$y_i\partial_i^2\partial_J F = r(i,J;y)F + \frac{1}{2}\sum_{k\in J}\frac{1}{y_i-y_k}(y_k\partial_k^2\partial_{J\setminus\{k\}})F. \tag{28}$$

The use of this recursive expression yields an efficient numerical evaluation of the matrices of the Pfaffian system. We keep numerical values of $\partial_k^2\partial_{J\setminus\{k\}}F$ in a table and use them to evaluate $\partial_i^2\partial_J F$ and keep it in the table, again.

We can also apply the recursion to the last term on the right-hand side. The resulting expression for $y_i\partial_i^2\partial_J F$ is given as

$$\begin{aligned}
y_i\partial_i^2\partial_J F = r(i,J;y)F &+ \frac{1}{2}\sum_{k_1\in J}\frac{1}{y_i-y_{k_1}}r(k_1,J\setminus\{k\};y)F \\
&+ \frac{1}{4}\sum_{\substack{k_1,k_2\in J\\ k_1,k_2:\text{distinct}}}\frac{1}{(y_i-y_{k_1})(y_{k_1}-y_{k_2})}r(k_2,J\setminus\{k_1,k_2\};y)F \\
&+ \frac{1}{8}\sum_{\substack{k_1,k_2,k_3\in J\\ k_1,k_2,k_3:\text{distinct}}}\frac{1}{(y_i-y_{k_1})(y_{k_1}-y_{k_2})(y_{k_2}-y_{k_3})}r(k_3,J\setminus\{k_1,k_2,k_3\};y)F + \dots \\
&+ \frac{1}{2^{|J|}}\sum_{\substack{k_1,\dots,k_{|J|}\in J\\ k_1,\dots,k_{|J|}:\text{distinct}}}\frac{1}{(y_i-y_{k_1})(y_{k_1}-y_{k_2})\dots(y_{k_{|J|-1}}-y_{k_{|J|}})}r(k_{|J|},\emptyset;y)F. \qquad (29)
\end{aligned}$$

Now in (4) we write $\Sigma^{-1}/2 = \beta = (\beta_1,\dots,\beta_m)$, where $\beta_1,\dots,\beta_m$ are distinct, and define a 2^m-dimensional vector valued function $\vec{G}$ in a scalar x by

$$\vec{G}(x) = \exp(-x\sum_{i=1}^m \beta_i)x^{mn/2}\vec{F}(\beta x).$$

Then $\vec{G}$ satisfies the ODE

$$\frac{d\vec{G}}{dx} = \left(-(\sum_{i=1}^m \beta_i)I_{2^m} + \frac{mn}{2x}I_{2^m} + \sum_{i=1}^m P_i(\beta x)\beta_i\right)\vec{G}, \qquad (30)$$

where I_{2^m} is the $2^m\times 2^m$ identity matrix. We denote the right-hand side as $P_\beta\vec{G}$. We now prove the following theorem, which is important for guaranteeing stability of ODE at $x=+\infty$.

Theorem 3. *As $x\to\infty$*

$$P_\beta = A_0 + O(1/x),$$

where A_0 only depends on β and the 2^m eigenvalues of A_0 are given as $-e_1\beta_1-\dots-e_m\beta_m$, where $(e_1,\dots,e_m)\in\{0,1\}^m$.

Proof. Note that $y_1=\beta_1x,\dots,y_m=\beta_mx=O(x)$. Divide (29) by $y_i=\beta_ix$. Then on the right-hand side of (29), the only constant order term is ∂_I in $r(i,J;y)$. Now

$$\begin{aligned}
\frac{d}{dx}\partial_I F(\beta x) &= \sum_{i=1}^m \beta_i\partial_i\partial_I F(\beta x) \\
&= \sum_{i\in I}\beta_i\partial_i^2\partial_{I\setminus\{i\}}F(\beta x) + \sum_{i\notin I}\beta_i\partial_{I\cup\{i\}}F(\beta x) \\
&= \big(\sum_{i\in I}\beta_i\big)\partial_I F(\beta x) + O(1/x) + \sum_{i\notin I}\beta_i\partial_{I\cup\{i\}}F(\beta x).
\end{aligned}$$

This implies that the I-th diagonal element of A_0 is given

$$-\sum_{i=1}^{m}\beta_i + \sum_{i\in I}\beta_i = -\sum_{i\notin I}\beta_i.$$

Furthermore the $(I, I\cup\{i\})$-element of A_0 is β_i. Other elements of A_0 are zeros. Hence A_0 is an upper triangular matrix with diagonal elements $-\sum_{i\notin I}\beta_i$, $I\subset[m]$. The theorem holds because the diagonal elements of an upper triangular matrix are its eigenvalues. $\square$

5 Some results of symbolic computation

In this section we present some results on symbolic computation for the initial values (cf. (17)) and the restriction for diagonal regions (cf. Section 3.2). We omit writing down the fully expanded form of (29), since the recursive formula (28) can be directly used in our implementation of holonomic gradient method.

5.1 Initial values

Initial values for our holonomic gradient method can be obtained by expressing ${}_1F_1$ in terms of monomial symmetric polynomials as in (16). We denote the relation between the zonal polynomials and the monomial symmetric polynomials in (15) as

$$\mathcal{C}_\kappa(Y) = \sum_{\lambda\trianglelefteq\kappa} c_{\kappa,\lambda}\mathcal{M}_\lambda(Y),$$

where $\lambda\trianglelefteq\kappa$ means that λ is dominated by κ, i.e. $\sum_{i=1}^{s}\lambda_i \le \sum_{i=1}^{s}\kappa_i$ for all s. Then ${}_1F_1$ is expressed as

$${}_1F_1(a;c;Y) = \sum_{k=0}^{\infty}\sum_{\lambda\vdash k} q_\lambda(a,c)\mathcal{M}_\lambda(Y), \quad q_\lambda(a,c) = \sum_{\kappa\vdash k,\kappa\trianglerighteq\lambda}\frac{(a)_\kappa c_{\kappa,\lambda}}{(c)_\kappa k!}.$$

A recurrence relation for $c_{\kappa,\lambda}$'s is given by James [9] (see also (14) in Section 7.2.1 of Muirhead [18] and Section 4.5.4 of Takemura [30]), which can be used to compute $q_\kappa(a,c)$. However James' recurrence relation works for each $\mathcal{C}_\kappa$ separately. Recently Koev and Edelman [10] gave a much improved algorithm based on recursive relations among the values of zonal polynomials for m variables and $m-1$ variables. For our implementation of holonomic gradient method, we adapted Koev-Edelman's recurrence relation also for derivatives of ${}_1F_1$ to evaluate the initial values.

Close to the origin, we can use rough initial values given by the linear approximation as in (18). Then we only need $\kappa=(k_1,\dots,k_l)$ such that $k_1=\cdots=k_l=1$ or $k_1=2$, $k_1=\cdots=k_l=1$. Some $q_\lambda(a,c)$'s for small λ's are as follows.

$$q_\varnothing = 1,\quad q_{(1)}=\frac{a}{c},\quad q_{(2)}=\frac{(a)_2}{2(c)_2},\quad q_{(1,1)}=\frac{(a)_2}{3(c)_2}+\frac{2(a)_{(1,1)}}{3(c)_{(1,1)}},\quad q_{(2,1)}=\frac{(a)_3}{10(c)_3}+\frac{2(a)_{(2,1)}}{5(c)_{(2,1)}},$$

$$q_{(1,1,1)}=\frac{(a)_3}{15(c)_3}+\frac{3(a)_{(2,1)}}{5(c)_{(2,1)}}+\frac{(a)_{(1,1,1)}}{3(c)_{(1,1,1)}},\ q_{(2,1,1)}=\frac{(a)_4}{70(c)_4}+\frac{4(a)_{(2,2)}}{45(c)_{(2,2)}}+\frac{11(a)_{(3,1)}}{63(c)_{(3,1)}}+\frac{2(a)_{(2,1,1)}}{9(c)_{(2,1,1)}},$$

where $\varnothing$ stands for the unique partition of zero. Write $(1^k) = (1, \ldots, 1), (2, 1^{k-2}) = (2, 1, \ldots, 1)$, which are partitions of k. Given the above quantities, the linear approximation of $\partial_1 \ldots \partial_l F(Y)$, $0 \le l \le m$, for $Y = (y_1, \ldots, y_m)$ close to the origin is expressed as

$$\partial_1 \ldots \partial_l F(Y) \doteq q_{(1^l)}(a, c) + 2q_{(2,1^{l-1})}(a, c)(y_1 + \cdots + y_l) + q_{(1^{l+1})}(a, c)(y_{l+1} + \cdots + y_m), \quad (31)$$

where for $l = 0$ the second term on the right-hand side is zero and for $l = m$ the third term is zero. We found that initial values by (31) are practical enough for $m \le 5$.

In fact, by Lemma 1 in Section 4.5.2 of [30] and by Proposition 7.3 of [25], $q_{(1^k)}(a, c)$ and $q_{(2,1^{k-2})}(a, c)$ are explicitly written as follows:

$$q_{(1^k)}(a, c) = 2^k k! \sum_{\kappa \vdash k} \frac{\prod_{1 \le i < j \le l(\kappa)} (2k_i - 2k_j - i + j)}{\prod_{i=1}^{l(\kappa)} (2k_i + l(\kappa) - i)!} \frac{(a)_\kappa}{(c)_\kappa},$$

$$q_{(2,1^{k-2})}(a, c) = 2^k (k-2)! \sum_{\kappa \vdash k} \frac{\prod_{1 \le i < j \le l(\kappa)} (2k_i - 2k_j - i + j)}{\prod_{i=1}^{l(\kappa)} (2k_i + l(\kappa) - i)!} \left(\binom{k}{2} + \sum_{i=1}^{l(\kappa)} k_i (k_i - i) \right) \frac{(a)_\kappa}{(c)_\kappa},$$

where $l(\kappa)$ is the length (number of non-zero parts) of $\kappa = (k_1, \ldots, k_{l(\kappa)})$.

For larger values of m we need higher order terms for initial values. For two partitions μ, λ, we write $\mu \subset \lambda$ to denote $\mu_i \le \lambda_i$ for all i. For two partitions κ, ν, we denote by $\kappa \uplus \nu$ the concatenation of κ and ν obtained from $(\kappa_1, \nu_1, \kappa_2, \nu_2, \ldots)$ by sorting. Consider a rectangular partition $\tau = (t, \ldots, t) = (t^l) \vdash tl$. For $\tau = (t^l)$ and $\lambda \supset \tau$ we define

$$I(\lambda, \tau) = \{(\kappa, \nu) \mid \kappa \uplus \nu = \lambda,\ \tau \subset \kappa,\ \kappa_{l+1} = 0\}.$$

Consider $\partial^\mu \mathcal{M}_\lambda(Y) = \partial_1^{\mu_1} \partial_2^{\mu_2} \cdots \partial_m^{\mu_m} \mathcal{M}_\lambda(Y)$. Note that $\partial^\mu \mathcal{M}_\lambda(Y) = 0$, if $\mu \not\subset \lambda$. For a rectangular $\tau = (t^l)$ we can calculate $\partial^\tau \mathcal{M}_\lambda(Y)$ by the following lemma.

Lemma 2. *For $\tau = (t^l) \subset \lambda$*

$$\partial^\tau \mathcal{M}_\lambda(Y) = \sum_{(\kappa, \nu) \in I(\lambda, \tau)} \frac{\kappa!}{(\kappa - \tau)!} \mathcal{M}_{\kappa - \tau}(y_1, \ldots, y_l) \mathcal{M}_\nu(y_{l+1}, \ldots, y_m),$$

where $\gamma! = \prod_i (\gamma_i!)$ for a partition γ, and $\kappa - \tau = (\kappa_1 - t, \ldots, \kappa_l - t)$.

Proof is straightforward and omitted. Using this lemma we have the following proposition.

Proposition 1. *For a rectangular partition $\tau = (t^l)$,*

$$\partial^\tau {}_1F_1(a; c; Y) = \sum_{k=tl}^{\infty} \sum_{\substack{\lambda \vdash k, \\ \tau \subset \lambda}} q_\lambda(a, c) \sum_{(\gamma, \nu) \in I(\lambda, \tau)} \frac{\gamma!}{(\gamma - \tau)!} \mathcal{M}_{\gamma - \tau}(y_1, \ldots, y_l) \mathcal{M}_\nu(y_{l+1}, \ldots, y_m). \quad (32)$$

For our initial values we only need to consider $\tau = (1^l)$. We obtain (31) if we only look at linear terms on the right-hand side of (32). Note that since ${}_1F_1(a; c; Y)$ is a symmetric function in $y_1, \ldots, y_m$, other derivatives are obtained by permutation of $y_1, \ldots, y_m$.

Although (32) only gives derivative with respect to a rectangular partition $\tau = (t^l)$, we can obtain other derivatives $\partial_1^{\mu_1} \ldots \partial_l^{\mu_h} {}_1F_1(a; c; Y)$, $\mu_1 \geq \cdots \geq \mu_h$, by repeated application of (32) for different values of l's.

5.2 Restriction to diagonal regions

As mentioned at the end of Section 3.2, the tedious operation involving the l'Hôpital rule for the diagonal region can be performed by the restriction algorithm for holonomic ideals. The following program in `Risa/Asir` for $m = 2$

```
import("names.rr"); import("nk_restriction.rr");
G1=y1*dy1^2 + (c-y1)*dy1+(1/2)*(y2/(y1-y2))*(dy1-dy2)-a;  G1=red((y1-y2)*G1);
G2=base_replace(G1,[[y1,y2],[y2,y1],[dy1,dy2],[dy2,dy1]]);
F=base_replace([G1,G2],[[y1,y],[y2,y+z2],[dy1,dy-dz2],[dy2,dz2]]);
A=nk_restriction.restriction_ideal(F,[z2,y],[dz2,dy],[1,0] | param=[a,c]);
```

produces the output

```
-y^2*dy^3+(3*y^2+(-3*c+1)*y)*dy^2+(-2*y^2+(4*a+4*c-2)*y-2*c^2+2*c)*dy-4*a*y+(4*c-4)*a
```

This is the same as (24). Adapting the above program for $m = 3$, we obtain

$$
\begin{aligned}
&y^3 f''''(y) + (-6y^3 + (6c-4)y^2) f'''(y) \\
&\quad + (11y^3 + (-10a - 22c + 18)y^2 + (11c^2 - 17c + 4)y) f''(y) \\
&\quad + (-6y^3 + (30a + 18c - 18)y^2 + ((-30c + 34)a - 18c^2 + 34c - 12)y \\
&\qquad\qquad + 6c^3 - 16c^2 + 10c) f'(y) \\
&\quad + (-18ay^2 + (9a^2 + (36c - 51)a)y + (-18c^2 + 48c - 30)a) f(y) = 0.
\end{aligned}
$$

For $m = 4$, we found that the computation by `Risa/Asir` takes too much time and memory.

We conjecture that the ideal I generated by $\prod_{j \neq i}(y_i - y_j) g_i$, $i = 1, \ldots, m$ in the Weyl algebra D_m is an holonomic ideal. In fact, the conjecture can be checked for small dimensions m on a computer. See the Appendix A. If I is a holonomic ideal, then

$$
J = (I + (y_1 - y_2)D_m + (y_1 - y_3)D_m + \cdots + (y_1 - y_m)D_m) \cap \mathbb{C}\langle y_1, \partial_{y_1} \rangle
$$

is not 0 and is an holonomic ideal in D_1 by the theorem of Bernstein (see, e.g., the Chapter 5 of [23]). The generators of J is ordinary differential equations for the function restricted to the diagonal $y_1 = \cdots = y_m$. Thus, the holonomicity implies the existence of the diagonal ordinary differential equation. The ideal J can be obtained by Oaku's algorithm ([20]) based on Gröbner bases and the `Risa/Asir` package `nk_restriction` uses this algorithm.

6 Numerical experiments

We implemented the holonomic gradient method in a straightforward manner. Our source programs in the language C are available from
`http://www.math.kobe-u.ac.jp/OpenXM/Math/1F1`.

The updating step of the holonomic gradient method was implemented using the recursive relation (28) for a general dimension. For initial values we adapted Koev and Edelman [10] for derivatives of ${}_1F_1$ as discussed in Section 5.1.

The accuracy of the holonomic gradient method can be simply checked by looking at the numerical convergence $\Pr[\ell_1 < x] \to 1$ as $x \to \infty$. This is because the initial values are evaluated at small $x > 0$ and $\Pr[\ell_1 < x]$ at large x is obtained after many updating steps. This is another advantage of our method.

Also we can use the following simple bounds for the upper tail probability for the purpose of checking. Let $\Pr[\ell_1 < x|\Sigma]$ denote the probability under the covariance matrix Σ. Consider $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, $\sigma_1^2 \geq \cdots \geq \sigma_m^2$. Then by standard stochastic ordering consideration, we have

$$\begin{aligned} \Pr[\ell_1 < x|\,\mathrm{diag}(\sigma_1^2, \ldots, \sigma_1^2)] &\leq \Pr[\ell_1 < x|\,\mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2)] \\ &\leq \Pr[\ell_1 < x|\,\mathrm{diag}(\sigma_1^2, 0, \ldots, 0)]. \end{aligned} \tag{33}$$

The upper bound coincides with the cumulative probability of chi-square distribution with n degrees of freedom (cf., [28],[31]). Accurate approximation for the lower bound $\Pr[\ell_1 < x|\sigma_1^2 I_m]$ is given by the tube method ([11], [12]).

We first consider the case $m = 5, n = 7$, $\Sigma^{-1}/2 = \beta = (1, 2, 3, 4, 5)$. For $x = 20$, the two bounds in (33) are given as 0.9996034 and 0.9999987. With the initial value of $x_0 = 0.01$ and step size 0.0001 we obtained

$$\Pr[\ell_1 < 20] = 0.999972.$$

The cumulative distribution function for this case is plotted on the left part of Figure 1.

Next we consider the case $m = 10, n = 12$, $\beta = (1, 2, \ldots, 10)$. For $x = 30$, the bounds are $(0.99866943, 0.99999998)$. For generation of initial values it takes about 20 seconds for approximating ${}_1F_1$ and its derivatives up to the degree 20 with an Intel Core i7 CPU. With the initial value of $x_0 = 0.2$ and step size 0.001, we obtain

$$\Pr[\ell_1 < 30] = 0.999545$$

in about 75 seconds. The cumulative distribution function for this case is plotted on the right part of Figure 1. We see that enough accuracy is obtained even for $m = 10$ within practical amount of time.

The complexity of numerically solving the ODE for $\vec{G}$ (30) is

$$O(m2^m) \times (\text{steps of the Runge-Kutta method with a prescribed precision}).$$

In fact, since the matrix $P_i(\beta x)$ has sparsity, each vector $P_i(\beta x)\vec{G}(x)$, which has 2^m elements, can be evaluated in $O(2^m)$ steps at x from the values of $\vec{G}(x)$ by utilizing (28).

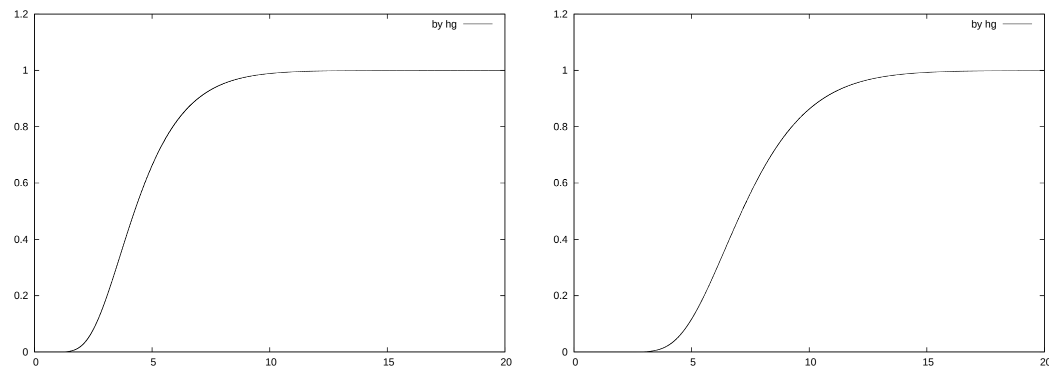

Figure 1: Cumulative distributions for $m = 5$ and $m = 10$

7 Discussion of open problems

The holonomic gradient method ([19]) gives a general algorithm for obtaining the partial differential equations satisfied by parametrized definite integrals such as the normalizing constant of a family of probability distributions. In fact, in [19] and Sei et al. [24] we used the holonomic gradient method for deriving the partial differential equations of the normalizing constants and for maximum likelihood estimation for distributions in directional statistics. For the case of ${}_1F_1$, the partial differential equations were already derived by Muirhead [16] more than 40 years ago. Our use of those partial differential equations for numerical evaluation of ${}_1F_1$ is very straightforward as discussed in Section 3 for the two dimensional case. Yet, from the viewpoint of holonomic functions, the partial differential equations of Muirhead [16] present many interesting open problems.

One important question is to obtain the ordinary and partial differential equations for the diagonal case as discussed in Section 3.2 for the case of $m = 2$. For a general dimension $m > 2$, it is desirable to be able to handle various patterns of diagonalization, such as the two-block diagonalization $y_1 = \cdots = y_l > y_{l+1} = \cdots = y_m$. A direct "by hand" calculation using the l'Hôpital rule becomes quickly infeasible when we increase m. Also the use of the restriction algorithm for holonomic ideals is limited by computational complexity. It is in fact a very heavy algorithm. Currently the `nk_restriction` routine of `Risa/Asir` in Section 5.2 takes too much time for $m \geq 4$. One possibility is to follow the approach in Muirhead [16] and Sugiyama et al. [29], where differentiation with respect to elementary symmetric functions of the roots of Y are considered. As discussed in Section 5.2, we conjecture that the ideal I generated by $\prod_{j \neq i}(y_i - y_j)g_i$, $i = 1, \ldots, m$ in the Weyl algebra D_m is an holonomic ideal. Holonomicity guarantees the existence of partial differential equations for diagonal regions.

Another question is to consider the asymptotics for $\Pr[\ell_1 \geq x] = 1 - \Pr[\ell_1 < x]$ as $x \to \infty$. As mentioned in the previous section, this tail probability can be approximated by the tube method ([11], [12]). One theoretical problem in applying the tube method is that only the approximation for the tail probability itself has been justified and the

justification of its derivatives has to be proved. However it is obvious that the current approach of taking the initial values close to the origin causes difficulty in precision for the extreme upper tail probability, in the case we want to evaluate the small probability $\Pr[\ell_1 \geq x]$. Hence it is desirable to be able to use tube formula approximation as the initial values at $x = \infty$.

From computational viewpoint, our holonomic gradient method has exponential complexity in the dimension m. We need to keep the 2^m-dimensional numerical vector $\vec{F}$ in memory at each step of the iteration. For $m = 20$, the dimension of the vector is about one million. Hence we do not expect that the current implementation of the holonomic gradient method works for $m = 20$. It might be possible to improve our current implementation by fully exploiting the fact that ${}_1F_1$ is a symmetric function in Y.

A Holonomicity for dimension two

In the theory of holonomic functions, the holonomicity of the left ideal generated by the set of partial differential operators is an important question. In fact, the existence of the ordinary differential equation with polynomial coefficients for the function restricted to the diagonal region follows from the holonomicity. Holonomicity of the ideal generated by g_1, g_2 in the two-dimensional case can be verified by Gröbner basis computation. Here we present this result. As to a general introduction to holonomic ideals and Gröbner bases, we refer to the Chapter 1 of Saito et al. [23].

Note that the holonomicity on the non-diagonal region $\mathcal{X}$ follows from Theorem 2 because the zero set of $y_i\xi_i^2 = 0$, $i = 1, \ldots, m$ contains the characteristic variety on $\mathcal{X}$. The holonomicity on $\mathcal{X}$ can also be proved by an analogous method with Ibukiyama et al. [7].

Let D_2 be the second Weyl algebra. For $P = \sum_{k=0}^{d} \sum_{\alpha_1+\alpha_2=k} f_{\alpha_1,\alpha_2}(y_1, y_2)\partial_1^{\alpha_1}\partial_2^{\alpha_2} \in D_2$, we define $\mathrm{in}(P)$ by

$$\mathrm{in}(P) = \sum_{\alpha_1+\alpha_2=d} f_{\alpha_1,\alpha_2}(y_1, y_2)\xi_1^{\alpha_1}\xi_2^{\alpha_2} \in \mathbb{C}[y_1, y_2, \xi_1, \xi_2],$$

where we assume that $f_{\alpha_1,\alpha_2}(y_1, y_2) \in \mathbb{C}[y_1, y_2]$ and that $f_{\alpha_1,\alpha_2}(y_1, y_2) \neq 0$ for some α_1, α_2 with $\alpha_1 + \alpha_2 = d$. For a left ideal I of D_2, the characteristic variety $\mathrm{ch}(I)$ is defined by

$$\mathrm{ch}(I) = \{(y_1, y_2, \xi_1, \xi_2) \in \mathbb{C}^{2\cdot 2} \mid \forall P \in I, \mathrm{in}(P)(y_1, y_2, \xi_1, \xi_2) = 0\}.$$

It is a basic fact that the dimension of the characteristic variety $\mathrm{ch}(I)$ of the proper left ideal I of D_2 is greater than or equal to 2. A left ideal I of D_2 is called *holonomic* if the dimension of the characteristic variety $\mathrm{ch}(I)$ equals 2.

Let $P_1 = (y_1 - y_2)g_1$, $P_2 = (y_2 - y_1)g_2$ and let I be the ideal of D_2 generated by P_1 and P_2. We will show that I is holonomic. Let $S = y_2\partial_2^2 P_1 + y_1\partial_1^2 P_2 + c(\partial_2 P_1 + \partial_1 P_2) \in I$. By

direct calculation we have

$$
\begin{aligned}
S =&(y_1^2y_2 - y_1y_2^2 + \frac{y_1^2}{2} - 2y_1y_2)\partial_1^2\partial_2 + (-y_1^2y_2 + y_1y_2^2 - 2y_1y_2 + \frac{y_2^2}{2})\partial_1\partial_2^2 - \frac{y_1^2}{2}\partial_1^3 - \frac{y_2^2}{2}\partial_2^3 \\
&+ (ay_1^2 - ay_1y_2 - \frac{3cy_1}{2} - y_1)\partial_1^2 + (-y_2 + 2ay_2 - ay_1y_2 + ay_2^2 - \frac{3cy_2}{2})\partial_2^2 \\
&+ (-cy_1^2 + 2cy_1y_2 - cy_2^2 + 4y_1y_2 - \frac{3cy_1}{2} - \frac{3cy_2}{2} + y_1 + y_2)\partial_1\partial_2 \\
&+ (acy_1 - acy_2 + 2ay_1 + cy_1 - c^2)\partial_1 + (-c^2 - acy_1 + acy_2 + cy_2)\partial_2 + 2ac.
\end{aligned}
$$

Hence

$$
\mathrm{in}(S) =(y_1^2y_2 - y_1y_2^2 + \frac{y_1^2}{2} - 2y_1y_2)\xi_1^2\xi_2 + (-y_1^2y_2 + y_1y_2^2 - 2y_1y_2 + \frac{y_2^2}{2})\xi_1\xi_2^2 - \frac{y_1^2}{2}\xi_1^3 - \frac{y_2^2}{2}\xi_2^3.
$$

Consider the ideal J of $\mathbb{C}[y_1, y_2, \xi_1, \xi_2]$ generated by $\mathrm{in}(P_1)$, $\mathrm{in}(P_2)$ and $\mathrm{in}(S)$. The following is a Gröbner base of J with respect to the graded reverse lexicographic order with $y_1 > y_2 > \xi_1 > \xi_2$:

$$
\begin{gathered}
\{y_2^2\xi_1^3\xi_2^2 + 3y_2^2\xi_1^2\xi_2^3 + 3y_2^2\xi_1\xi_2^4 + y_2^2\xi_2^5, y_1y_2\xi_1^3 + 3y_1y_2\xi_1^2\xi_2 + 3y_2^2\xi_1\xi_2^2 + y_2^2\xi_2^3, \\
y_1^2\xi_1^2 - y_1y_2\xi_1^2, y_1y_2\xi_2^2 - y_2^2\xi_2^2\}.
\end{gathered}
$$

Thus the Krull dimension of J is 2. Since $\{P_1, P_2, S\} \subset I$, the characteristic variety $\mathrm{ch}(I)$ of I is contained in the zero set of J. This implies that the dimension of $\mathrm{ch}(I)$ does not exceed 2 and hence the dimension of $\mathrm{ch}(I)$ is equal to 2. Therefore I is holonomic for $m = 2$.

B R source program for dimension two

The following program in data analysis system R implements holonomic gradient method for $m = 2$ of Section 3.1 based on deSolve add-on package for R.

```
library(deSolve)
m <- 2    # dimension
n <- 3    # degrees of freedom
x <- 4.31600  #  specify x.   We evaluate  Pr( l1 < x )
b1 <- 1; b2 <- 2 #  (b1,b2) =  (1/2) diag(Sigma^{-1})
a <- (m+1)/2; c <- (n+m+1)/2;  totalsteps <- 10000; stepsize <- x/totalsteps

# h's
h2000 <- function(y1,y2) a/y1
h2010 <- function(y1,y2) -(c-y1)/y1 - y2/(2*y1*(y1-y2))
h2001 <- function(y1,y2) y2/(2*y1*(y1-y2))
h1200 <- function(y1,y2) a/(2*y2*(y2-y1))
h1210 <- function(y1,y2) 3/(4*(y2-y1)^2) + a/y2 - (c-y1)/(2*y2*(y2-y1))
h1201 <- function(y1,y2) -3/(4*(y2-y1)^2)
h1211 <- function(y1,y2) -(c-y2)/y2 - y1/(2*y2*(y2-y1))

#initial values
x1 <- b1*stepsize; x2 <- b2*stepsize
fi <- c(a/c, (a*(a+1))/(c*(c+1)), (a*(a+1))/(3*c*(c+1)) + (2*a*(a-1/2))/(3*c*(c-1/2)),
```

```
       (a*(a+1)*(a+2))/(5*c*(c+1)*(c+2)) + (4*a*(a+1)*(a-1/2))/(5*c*(c+1)*(c-2/1)))
fi <- c(1+(x1+x2)*fi[1], fi[1]+x1*fi[2]+x2*fi[3], fi[1]+x2*fi[2]+x1*fi[3], fi[3]+(x1+x2)*fi[4])

# gradient
f11m2 <- function(y,fv,parm){y1 <- y*b1; y2 <- y*b2;
      list(c(
        b1*fv[2] + b2*fv[3],
        b1*(fv[1]*h2000(y1,y2)+fv[2]*h2010(y1,y2)+fv[3]*h2001(y1,y2)) + b2*fv[4],
        b2*(fv[1]*h2000(y2,y1)+fv[2]*h2001(y2,y1)+fv[3]*h2010(y2,y1)) + b1*fv[4],
         b1*(fv[1]*h1200(y2,y1)+fv[2]*h1201(y2,y1)+fv[3]*h1210(y2,y1)+fv[4]*h1211(y2,y1))
        +b2*(fv[1]*h1200(y1,y2)+fv[2]*h1210(y1,y2)+fv[3]*h1201(y1,y2)+fv[4]*h1211(y1,y2)))
)}

output <- ode(fi,func=f11m2,(1:totalsteps)*x/totalsteps)
prob0 <- ((b1*b2)^(n/2)*gamma(a)*gamma(a-1/2))/(gamma(c)*gamma(c-1/2)) * x^(n*m/2) * exp(-x*(b1+b2))
cat("x=",x, "prob=", output[totalsteps,2]*prob0,"\n")
```

References

[1] Uri M. Ascher and Linda R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. ISBN 0-89871-412-5.

[2] Ronald W. Butler and Andrew T. A. Wood. Laplace approximations for hypergeometric functions with matrix argument. *Ann. Statist.*, 30(4):1155–1177, 2002. ISSN 0090-5364.

[3] A. G. Constantine. Some non-central distribution problems in multivariate analysis. *Ann. Math. Statist.*, 34:1270–1285, 1963. ISSN 0003-4851.

[4] Kenneth I. Gross and Donald St. P. Richards. Special functions of matrix argument. I. Algebraic induction, zonal polynomials, and hypergeometric functions. *Trans. Amer. Math. Soc.*, 301(2):781–811, 1987. ISSN 0002-9947.

[5] Hiroki Hashiguchi and Naoto Niki. Numerical computation on distributions of the largest and the smallest latent roots of the Wishart matrix. *J. Japanese Soc. Comput. Statist.*, 19(1):45–56, 2006. ISSN 0915-2350.

[6] Carl S. Herz. Bessel functions of matrix argument. *Ann. of Math. (2)*, 61:474–523, 1955. ISSN 0003-486X.

[7] Tomoyoshi Ibukiyama, Takako Kuzumaki, and Hiroyuki Ochiai. Holonomic systems of Gegenbauer type polynomials of matrix arguments related with Siegel modular forms. *Journal of the Mathematical Society of Japan*. In Press.

[8] Alan T. James. Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.*, 35:475–501, 1964. ISSN 0003-4851.

[9] Alan T. James. Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist*, 39:1711–1718, 1968. ISSN 0003-4851.

[10] Plamen Koev and Alan Edelman. The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comp.*, 75(254):833–846, 2006. ISSN 0025-5718.

[11] Satoshi Kuriki and Akimichi Takemura. Tail probabilities of the maxima of multilinear forms and their applications. *Ann. Statist.*, 29(2):328–371, 2001. ISSN 0090-5364.

[12] Satoshi Kuriki and Akimichi Takemura. Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant random matrix. *J. Statist. Plann. Inference*, 138(11):3357–3378, 2008. ISSN 0378-3758.

[13] Vadim B. Kuznetsov and Siddhartha Sahi, editors. *Jack, Hall-Littlewood and Macdonald Polynomials*, volume 417 of *Contemporary Mathematics*, Providence, RI, 2006. American Mathematical Society. ISBN 0-8218-3683-8.

[14] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. ISBN 0-19-853489-2. With contributions by A. Zelevinsky, Oxford Science Publications.

[15] A. M. Mathai, Serge B. Provost, and Takesi Hayakawa. *Bilinear Forms and Zonal Polynomials*, volume 102 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1995. ISBN 0-387-94522-9.

[16] R. J. Muirhead. Systems of partial differential equations for hypergeometric functions of matrix argument. *Ann. Math. Statist.*, 41:991–1001, 1970. ISSN 0003-4851.

[17] Robb J. Muirhead. Latent roots and matrix variates: a review of some asymptotic results. *Ann. Statist.*, 6(1):5–33, 1978. ISSN 0090-5364.

[18] Robb J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons Inc., New York, 1982. ISBN 0-471-09442-0. Wiley Series in Probability and Mathematical Statistics.

[19] Hiromasa Nakayama, Kenta Nishiyama, Masayuki Noro, Katsuyoshi Ohara, Tomonari Sei, Nobuki Takayama, and Akimichi Takemura. Holonomic gradient descent and its application to the Fisher-Bingham integral. *Advances in Applied Mathematics*, 47:639–658, 2011. ISSN 0196-8858.

[20] Toshinori Oaku. Algorithms for b-functions, restrictions, and algebraic local cohomology groups of D-modules. *Adv. in Appl. Math.*, 19(1):61–105, 1997. ISSN 0196-8858.

[21] D. St. P. Richards. Functions of matrix argument. In *NIST Handbook of Mathematical Functions*, pages 767–774. U.S. Dept. Commerce, Washington, DC, 2010.

[22] RisaAsir developing team. Risa/asir, a computer algebra system. Available at http://www.math.kobe-u.ac.jp/Asir/.

[23] Mutsumi Saito, Bernd Sturmfels, and Nobuki Takayama. *Gröbner Deformations of Hypergeometric Differential Equations*, volume 6 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2000. ISBN 3-540-66065-8.

[24] Tomonari Sei, Hiroki Shibata, Akimichi Takemura, Katsuyoshi Ohara, and Nobuki Takayama. Properties and applications of Fisher distribution on the rotation group, 2011. arXiv:1110.0721v1.

[25] Richard P. Stanley. Some combinatorial properties of Jack symmetric functions. *Adv. Math.*, 77(1):76–115, 1989. ISSN 0001-8708.

[26] W. A. Stein et al. *Sage Mathematics Software (Version 4.7.2)*. The Sage Development Team, 2011. `http://www.sagemath.org`.

[27] Takakazu Sugiyama. On the distribution of the largest latent root of the covariance matrix. *Ann. Math. Statist*, 38:1148–1151, 1967. ISSN 0003-4851.

[28] Takakazu Sugiyama. Approximation for the distribution of the largest latent root of a Wishart matrix. *Austral. J. Statist.*, 14:17–24, 1972. ISSN 0004-9581.

[29] Takakazu Sugiyama, Yuuichi Takeda, and Masafumi Fukuda. On the numerical computation of confluent hypergeometric function with zonal polynomials of order 3. *J. Japanese Soc. Computat. Statist.*, 11(1):1–8, 1998. ISSN 0915-2350.

[30] Akimichi Takemura. *Zonal Polynomials*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 4. Institute of Mathematical Statistics, Hayward, CA, 1984. ISBN 0-940600-05-6.

[31] Akimichi Takemura and Yo Sheena. Distribution of eigenvalues and eigenvectors of Wishart matrix when the population eigenvalues are infinitely dispersed and its application to minimax estimation of covariance matrix. *J. Multivariate Anal.*, 94 (2):271–299, 2005. ISSN 0047-259X.

[32] Doron Zeilberger. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.*, 32(3):321–368, 1990. ISSN 0377-0427.